SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Carter, Kenneth C.
               He, Wei-Wu

    (ii) TITLE OF INVENTION: Human NK-3 Related Prostate Specific
         Gene-1

  (iii) NUMBER OF SEQUENCES: 30

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
        (B) STREET: 1100 NEW YORK AVE., NW, SUITE 600
        (C) CITY: WASHINGTON
        (D) STATE: DC
        (E) COUNTRY: USA
        (F) ZIP: 20005

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: US 09/105,470
        (B) FILING DATE: 26-JUN-1998
        (C) CLASSIFICATION:

   (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER: US 60/051,080
        (B) FILING DATE: 27-JUN-1997

  (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: STEFFE, ERIC K.
        (B) REGISTRATION NUMBER: 36,688
        (C) REFERENCE/DOCKET NUMBER: 1488.0790001

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: 301-309-8504
        (B) TELEFAX: 301-309-8439

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 705 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGCTCAGGG TTCCGGAGCC GCGGCCCGGG GAGGCGAAAG CGGAGGGGGC CGCGCCGCCG   60

```
ACCCCGTCCA AGCCGCTCAC GTCCTTCCTC ATCCAGGACA TCCTGCGGGA CGGCGCGCAG       120

CGGCAAGGCG GCCGCACGAG CAGCCAGAGA CAGCGCGACC CGGAGCCGGA GCCAGAGCCA       180

GAGCCAGAGG GAGGACGCAG CCGCGCCGGG GCGCAGAACG ACCAGCTGAG CACCGGGCCC       240

CGCGCCGCGC CGGAGGAGGC CGAGACGCTG GCAGAGACCG AGCCAGAAAG GCACTTGGGG       300

TCTTATCTGT TGGACTCTGA AAACACTTCA GGCGCCCTTC CAAGGCTTCC CCAAACCCCT       360

AAGCAGCCGC AGAAGCGCTC CCGAGCTGCC TTCTCCCACA CTCAGGTGAT CGAGTTGGAG       420

AGGAAGTTCA GCCATCAGAA GTACCTGTCG GCCCCTGAAC GGGCCCACCT GGCCAAGAAC       480

CTCAAGCTCA CGGAGACCCA AGTGAAGATA TGGTTCCAGA ACAGACGCTA TAAGACTAAG       540

CGAAAGCAGC TCTCCTCGGA GCTGGGAGAC TTGGAGAAGC ACTCCTCTTT GCCGGCCCTG       600

AAAGAGGAGG CCTTCTCCCG GGCCTCCCTG GTCTCCGTGT ATAACAGCTA TCCTTACTAC       660

CCATACCTGT ACTGCGTGGG CAGCTGGAGC CCAGCTTTTG GGTAA                       705
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 234 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Leu Arg Val Pro Glu Pro Arg Pro Gly Glu Ala Lys Ala Glu Gly
1               5                   10                  15

Ala Ala Pro Pro Thr Pro Ser Lys Pro Leu Thr Ser Phe Leu Ile Gln
            20                  25                  30

Asp Ile Leu Arg Asp Gly Ala Gln Arg Gln Gly Gly Arg Thr Ser Ser
        35                  40                  45

Gln Arg Gln Arg Asp Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Gly
    50                  55                  60

Gly Arg Ser Arg Ala Gly Ala Gln Asn Asp Gln Leu Ser Thr Gly Pro
65                  70                  75                  80

Arg Ala Ala Pro Glu Glu Ala Glu Thr Leu Ala Glu Thr Glu Pro Glu
                85                  90                  95

Arg His Leu Gly Ser Tyr Leu Leu Asp Ser Glu Asn Thr Ser Gly Ala
                100                 105                 110

Leu Pro Arg Leu Pro Gln Thr Pro Lys Gln Pro Gln Lys Arg Ser Arg
            115                 120                 125

Ala Ala Phe Ser His Thr Gln Val Ile Glu Leu Glu Arg Lys Phe Ser
    130                 135                 140
```

```
        His Gln Lys Tyr Leu Ser Ala Pro Glu Arg Ala His Leu Ala Lys Asn
        145             150                 155                 160

        Leu Lys Leu Thr Glu Thr Gln Val Lys Ile Trp Phe Gln Asn Arg Arg
                        165                 170                 175

        Tyr Lys Thr Lys Arg Lys Gln Leu Ser Ser Glu Leu Gly Asp Leu Glu
                    180                 185                 190

        Lys His Ser Ser Leu Pro Ala Leu Lys Glu Glu Ala Phe Ser Arg Ala
                195                 200                 205

        Ser Leu Val Ser Val Tyr Asn Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr
            210                 215                     220

        Cys Val Gly Ser Trp Ser Pro Ala Phe Gly
        225                 230
```

(2) INFORMATION FOR SEQ ID NO:3:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 705 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATGCTCAGGG TTCCGGAGCC GCGGCCCGGG GAGGCGAAAG CGGAGGGGGC CGCGCCGCCG      60

ACCCCGTCCA AGCCGCTCAC GTCCTTCCTC ATCCAGGACA TCCTGCGGGA CGGCGCGCAG     120

CGGCAAGGCG GCCGCACGAG CAGCCAGAGA CAGTGCGACC CGGAGCCGGA GCCAGAGCCA     180

GAGCCAGAGG GAGGACGCAG CCGCGCCGGG GCGCAGAACG ACCAGCTGAG CACCGGGCCC     240

CGCGCCGCGC CGGAGGAGGC CGAGACGCTG GCAGAGACCG AGCCAGAAAG GCACTTGGGG     300

TCTTATCTGT TGGACTCTGA AAACACTTCA GGCGCCCTTC CAAGGCTTCC CCAAACCCCT     360

AAGCAGCCGC AGAAGCGCTC CCGAGCTGCC TTCTCCCACA CTCAGGTGAT CGAGTTGGAG     420

AGGAAGTTCA GCCATCAGAA GTACCTGTCG GCCCCTGAAC GGGCCCACCT GGCCAAGAAC     480

CTCAAGCTCA CGGAGACCCA AGTGAAGATA TGGTTCCAGA ACAGACGCTA TAAGACTAAG     540

CGAAAGCAGC TCTCCTCGGA GCTGGGAGAC TTGGAGAAGC ACTCCTCTTT GCCGGCCCTG     600

AAAGAGGAGG CCTTCTCCCG GGCCTCCCTG GTCTCCGTGT ATAACAGCTA TCCTTACTAC     660

CCATACCTGT ACTGCGTGGG CAGCTGGAGC CCAGCTTTTG GGTAA                     705
```

(2) INFORMATION FOR SEQ ID NO:4:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 234 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

    Met Leu Arg Val Pro Glu Pro Arg Pro Gly Glu Ala Lys Ala Glu Gly
    1   5   10   15

    Ala Ala Pro Pro Thr Pro Ser Lys Pro Leu Thr Ser Phe Leu Ile Gln
                20   25   30

    Asp Ile Leu Arg Asp Gly Ala Gln Arg Gln Gly Gly Arg Thr Ser Ser
            35   40   45

    Gln Arg Gln Cys Asp Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Gly
        50   55   60

    Gly Arg Ser Arg Ala Gly Ala Gln Asn Asp Gln Leu Ser Thr Gly Pro
    65   70   75   80

    Arg Ala Ala Pro Glu Glu Ala Glu Thr Leu Ala Glu Thr Glu Pro Glu
                85   90   95

    Arg His Leu Gly Ser Tyr Leu Leu Asp Ser Glu Asn Thr Ser Gly Ala
                100   105   110

    Leu Pro Arg Leu Pro Gln Thr Pro Lys Gln Pro Gln Lys Arg Ser Arg
            115   120   125

    Ala Ala Phe Ser His Thr Gln Val Ile Glu Leu Glu Arg Lys Phe Ser
        130   135   140

    His Gln Lys Tyr Leu Ser Ala Pro Glu Arg Ala His Leu Ala Lys Asn
    145   150   155   160

    Leu Lys Leu Thr Glu Thr Gln Val Lys Ile Trp Phe Gln Asn Arg Arg
                165   170   175

    Tyr Lys Thr Lys Arg Lys Gln Leu Ser Ser Glu Leu Gly Asp Leu Glu
                180   185   190

    Lys His Ser Ser Leu Pro Ala Leu Lys Glu Glu Ala Phe Ser Arg Ala
            195   200   205

    Ser Leu Val Ser Val Tyr Asn Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr
        210   215   220

    Cys Val Gly Ser Trp Ser Pro Ala Phe Gly
    225   230

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 60 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Lys Lys Arg Ser Arg Ala Ala Phe Ser His Ala Gln Val Phe Glu Leu
1   5   10   15

Glu Arg Arg Phe Ala Gln Gln Arg Tyr Leu Ser Gly Pro Glu Arg Ser
                20   25   30

Glu Met Ala Lys Ser Leu Arg Leu Thr Glu Thr Gln Val Lys Ile Trp
        35   40   45

Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln
    50   55   60

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 60 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein.

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Lys Lys Arg Lys Arg Arg Val Leu Phe Thr Lys Ala Gln Thr Tyr Glu
1   5   10   15

Leu Glu Arg Arg Phe Arg Gln Gln Arg Tyr Leu Ser Ala Pro Glu Arg
                20   25   30

Glu His Leu Ala Ser Leu Ile Arg Leu Thr Pro Thr Gln Val Lys Ile
        35   40   45

Trp Phe Gln Asn His Arg Tyr Lys Thr Lys Arg Ala
    50   55   60

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 60 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Lys Arg Lys Pro Arg Val Leu Phe Ser Gln Ala Gln Val Leu Glu Leu
1   5   10   15

Glu Cys Arg Phe Arg Leu Lys Lys Tyr Leu Thr Gly Ala Glu Arg Glu

          Ile Ile Ala Gln Lys Leu Asn Leu Ser Ala Thr Gln Val Lys Ile Trp
              35   40   45

          Phe Gln Asn Arg Arg Tyr Lys Ser Lys Arg Gly Asp
              50   55   60

    (2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 3488 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AATTAACCCT CACTAAAGGG AACAAAAGCT GGAGCTCCAC CGCGGTGGCG GCCGCGTAAT   60

ACGACTCACT ATAGGGCGAA GAATTCGGAT CTATCAATCT GCATCCTTGT TTCAGAACCA   120

TTTGATGTAA GTTTCATAAA TCTTGTGCCT TTGCTCCTAC TTACTTCAGT GTTTATTTCC   180

TAAAAATATT CTCTTGTACA CTGACAGTAC AATGTGCAAT TTCAGTAAAT TTAACATTAA   240

TTCAATACTT CCATCATCGA CCTGACACTG AGACTCATGC CTGTAGTCCT GGCACTTTGA   300

GAGGCCAAGG CAGGAGGATC ACTTGAATCC AGGAAATCGA GGCTGCAGTG AGTTATGATG   360

GCATCACTGC ACTCCAGCCT GGGCGGCAGA GGGAGACCCT GTCCGTAAAA AACAGAAGAG   420

AAAAGACAAG GAAAGAAAAT ACTTCCATCA TCTCTGTTCC ACTTTCGTCT GTTGTCACGG   480

TACCGTCCAG TCCAGTCACA GTACCGGTTG GACCAATCTG GCTAACCCAT TGTTTAGCCA   540

ATGGGTTACA TGTTAACAGT TGGTAATCTG CAAAAAGAGT ATGCTGATGT TCTTTTGAAC   600

TACTTTTTTA AATGCAGTTT TTGCATTTGT CCCTGGCCTA AAACGCCTTC CATCCGTCTG   660

GAAACTTTTC AAAAGGATGG TATGTCATGT GTCTGGGGAG GAAGGAAAGT TAACAGGTTA   720

TTGCGGATAA AGGAACCACC AAAGAAAACC ACTTCTGCAA CGGGAAAAGG CTTTGGCAAA   780

GGTGTTTTCC TTCTTTCAGC CTGGGGTCTG GCTGCACCTA CTTGTCATGC CTCTTTGAGG   840

TCGTAGATAT TGCAGATCTG AGTTTGCACC ATCTCTCCCA GAGAGAGAGA GCACCCAGAA   900

CTCTCACGGT ACCGCGCGGC TGCAGTGACT GCGTGCTCAT CCCCTGTAAT TGGCTCTGAC   960

GGTCCTGAAG AGCTAACTGG ACTGTTTGTC TTGATCGTCC CATCCCCAGG AGCTTCTCTC   1020

TGCTGCGGGT GGGTTGGGGC AGAGGAGCCC CGCTTTGGGG TGCGCTCCTG GCCTGGGAAA   1080

ACGGCTCAGG GCGGAGGGAG GAGAGCTGGA GAAGGAGAGG AAATTGGGGA AGGAGAGGGA   1140

```
ATTGGGGAAG GAGAGGGAAC TGGGGAAGGA ATCCCCTAGG GAGGAGCGGA GCGGGGCAGT      1200

GCTCAGGGCT CGCAGATCGG CGGGGTCACC TGGGGCTCAG GGCGGCCAAT CCGCGGCGCG      1260

GCCCGTCCCG CGGCCAATGG GAGGGCGGCG CGGCCCGCTC CCCTGGGCTA TAAGCGAGCC      1320

GGGAGGCGGA AAGTGAAAGC GGTGCGGGCC GGGCGGGTGC ATTCAGGCCA AGGCGGGGCC      1380

GCCGGGATGC TCAGGGTTCC GGAGCCGCGG CCCGGGGAGG CGAAAGCGGA GGGGGCCGCG      1440

CCGCCGACCC CGTCCAAGCC GCTCACGTCC TTCCTCATCC AGGACATCCT GCGGGACGGC      1500

GCGCAGCGGC AAGGCGGCCG CACGAGCAGC CAGAGACAGC GCGACCCGGA GCCGGAGCCA      1560

GAGCCAGAGC CAGAGGGAGG ACGCAGCCGC GCCGGGGCGC AGAACGACCA GCTGAGCACC      1620

GGGCCCCGCG CCGCGCCGGA GGAGGCCGAG ACGCTGGCAG AGACCGAGCC AGGTAAGCGG      1680

CGAGGCCGGG GAAGGGGGGC AGCCCAAGGC GGACCCCCAG AGCTCGGGGT GCAGGGACGC      1740

GGGGCTCCGC GGCGACAGGC AGAGGGACCT TCCCGCCTCC GCAGCCACGC GCGCGCCCCC      1800

GGAATGAACC CTGAGCCCCA GCGTCAGGGC GGCGCAGGAT TCTGACACCG CAGGATTCGC      1860

CCGGTTCCGT GCCTTCCGTT CCCTGGGGCT CAGAAGCCGG CGCGACTGCA GCGCCACCGC      1920

CTTCCACCGT CCCAGGAGCG GATCCCGCCC CGCGCCACCC GCGATCGGCG CCAGCCCCCC      1980

GGTAGTTATG AGAANTAATA ATAACTTATT AACAGTGACA AAGCAGGGGT TGACCAGCAA      2040

AGCCTCCGTG TGCTTCCCAA TCCCGTGGGC AGTAAAGCGG TATATTCGGG GTTCCCTCCG      2100

GTGTCCAGGA GAGAGAGTCC ACTTATTTTC TTTCCTGTCA CTTCTGATGA GGCGACCGAA      2160

CGCCTCGTTT AGCGAAGAGG GAATTAAAGC CCAGAATGAG CCTGCCTCTG CGTCTCCAGT      2220

GGCACAAGCC CTCTCTTGCC CACCTGGATC CTAACACCGG ATGTCTTTTG GTCTGGCCTT      2280

CCCGGGTATC TTGTTCCACG GCATTTTCCC TGCCTCCCTC TCCCGCCTCT CCTCAGCACA      2340

CAGATCCAGA ATCCCCATAT AATTCTACTA GACAGTAGGG AGAAAGTTCA ACCACGAAAC      2400

GTCTCTAACT TTGGGTTCTT GATGATTCTT AGCAAATGAA TGCGTAATAA ACATATTTAC      2460

TCACTCTTCA CTCCGGAGAG CTCCTTAGTC ATGTGAAAAA AGTGAAATGT ATCCACGATG      2520

ACAGTGGGCT GTTTGTTCAC TCACTAAAGA GATAAGGGTG GATTGAATTC TCTTCTCTTC      2580

CCTGCTAACA TGTAACTTTT GTCTTCCCAT CCCTCCTTCC CCACTCTCCT TTCCAGAAAG      2640

GCACTTGGGG TCTTATCTGT TGGACTCTGA AAACACTTCA GGCGCCCTTC CAAGGCTTCC      2700

CCAAACCCCT AAGCAGCCGC AGAAGCGCTC CCGAGCTGCC TTCTCCCACA CTCAGGTGAT      2760

CGAGTTGGAG AGGAAGTTCA GCCATCAGAA GTACCTGTCG GCCCCTGAAC GGGCCCACCT      2820

GGCCAAGAAC CTCAAGCTCA CGGAGACCCA AGTGAAGATA TGGTTCCAGA ACAGACGCTA      2880

TAAGACTAAG CGAAAGCAGC TCTCCTCGGA GCTGGGAGAC TTGGAGAAGC ACTCCTCTTT      2940

GCCGGCCCTG AAAGAGGAGG CCTTCTCCCG GGCCTCCCTG GTCTCCGTGT ATAACAGCTA      3000
```

```
TCCTTACTAC CCATACCTGT ACTGCGTGGG CAGCTGGAGC CCAGCTTTTG GGTAATGCCA    3060

GCTCAGGTGA CAACCATTAT GATCAAAAAC TGCCTTCCCC AGGGTGTCTC TATGAAAAGC    3120

ACAAGGGGCC AAGGTCAGGG AGCAAGAGGT GTGCACACCA AAGCTATTGG AGATTTGCGT    3180

GGAAATCTCA GATTCTTCAC TGGTGAGACA ATGAAACAAC AGAGACAGTG AAAGTTTTAA    3240

TACCTAAGTC ATTCCTCCAG TGCATACTGT AGGTCATTTT TTTTGGTTCT GGCTACCTGT    3300

TTGAAGGGGA GAGAGGGAAA ATCAAGTGGT ATTTTCCAGC ACTTTGTATG ATTTTGGATG    3360

AGTTGTACAC CCAAGGATTC TGTTATGCAA CTCCATCCTC CTGTGTCACT GAATATCAAC    3420

TCTGAAAGAG CAAACCTAAC AGGAGAAAGG ACAACCAGGA TGAGGATGTC ACCAACTGAA    3480

TTAAACTC                                                             3488
```

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 124 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Leu Arg Val Ala Glu Pro Arg Glu Pro Arg Val Glu Ala Gly Gly
1               5                   10                  15

Arg Ser Pro Trp Ala Ala Pro Pro Thr Gln Ser Lys Arg Leu Thr Ser
            20                  25                  30

Phe Leu Ile Gln Asp Ile Leu Arg Asp Arg Ala Glu Arg His Gly Gly
        35                  40                  45

His Ser Gly Asn Pro Gln His Ser Pro Asp Pro Arg Arg Asp Ser Ala
    50                  55                  60

Pro Glu Pro Asp Lys Ala Gly Gly Arg Gly Val Ala Pro Glu Asp Pro
65                  70                  75                  80

Pro Ser Ile Arg His Ser Pro Ala Glu Thr Pro Thr Glu Pro Glu Ser
            85                  90                  95

Asp Ala His Phe Glu Thr Tyr Leu Leu Asp Cys Glu His Asn Pro Gly
            100                 105                 110

Asp Leu Ala Ser Ala Pro Gln Val Thr Lys Gln Pro
    115                 120
```

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 122 amino acids
       (B) TYPE: amino acid

```
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

    Met Leu Arg Val Pro Glu Pro Arg Pro Gly Glu Ala Lys Ala Glu Gly
    1               5                   10                  15

    Ala Ala Pro Pro Thr Pro Ser Lys Pro Leu Thr Ser Phe Leu Ile Asp
                20                  25                  30

    Ile Leu Arg Asp Gly Ala Gln Arg Gln Gly Gly Arg Thr Ser Ser Gln
            35                  40                  45

    Arg Gln Cys Asp Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Gly Gly
        50                  55                  60

    Arg Ser Arg Ala Gly Ala Gln Asn Asp Gln Leu Ser Thr Gly Pro Arg
    65                  70                  75                  80

    Ala Ala Pro Glu Glu Ala Glu Thr Leu Ala Glu Thr Glu Pro Glu Arg
                85                  90                  95

    His Leu Gly Ser Tyr Leu Leu Asp Ser Glu Asn Thr Ser Gly Ala Leu
                100                 105                 110

    Pro Arg Leu Pro Gln Thr Pro Lys Gln Pro
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 60 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
    Gln Lys Arg Ser Arg Ala Ala Phe Ser His Thr Gln Val Ile Glu Leu
    1               5                   10                  15

    Glu Arg Lys Phe Ser His Gln Lys Tyr Leu Ser Ala Pro Glu Arg Ala
                20                  25                  30

    His Leu Ala Lys Asn Leu Lys Leu Thr Glu Thr Gln Val Lys Ile Trp
            35                  40                  45

    Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln
        50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

   Gln Lys Arg Ser Arg Ala Ala Phe Ser His Thr Gln Val Ile Glu Leu
   1   5   10   15

   Glu Arg Lys Phe Ser His Gln Lys Tyr Leu Ser Ala Pro Glu Arg Ala
               20   25   30

   His Leu Ala Lys Asn Leu Lys Leu Thr Glu Thr Gln Val Lys Ile Trp
           35   40   45

   Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln
       50   55   60

(2) INFORMATION FOR SEQ ID NO:13:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 53 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

   Leu Ser Glu Asp Leu Gly Val Leu Glu Lys Asn Ser Pro Leu Ser Leu
   1   5   10   15

   Pro Ala Leu Lys Asp Asp Ser Leu Pro Ser Thr Ser Leu Val Ser Val
               20   25   30

   Tyr Thr Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr Cys Leu Gly Ser Trp
           35   40   45

   His Pro Ser Phe Trp
       50

(2) INFORMATION FOR SEQ ID NO:14:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 51 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
        Leu Ser Ser Glu Leu Gly Asp Leu Glu Lys His Ser Ser Leu Pro Ala
        1               5                   10                  15

        Leu Lys Glu Glu Ala Phe Ser Arg Ala Ser Leu Val Ser Val Tyr Asn
                    20                  25                  30

        Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr Cys Val Gly Ser Trp Ser Pro
                35                  40                  45

        Ala Phe Gly
            50
```

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGGGATCCA TGCTCAGGGT TCCGGAG   27

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCGAGCTTTT ACCCAAAAGC TGGGCT   26

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCGGGATCCC ATGCTCAGGG TTCCGGAG   28

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 26 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCGGATCCTT ACCCAAAAGC TGGGCT   26

(2) INFORMATION FOR SEQ ID NO:19:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 26 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCGGATCCTT ACCCAAAAGC TGGGCT   26

(2) INFORMATION FOR SEQ ID NO:20:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CACTGCCCAG TCAAGTGTTC TTGA   24

(2) INFORMATION FOR SEQ ID NO:21:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CACTGCCCAG TCACGTGTTC GTGA   24

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 15 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ACACTAATTG GAGGC   15

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 15 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACACTACTTG GAGGC   15

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 19 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CTCTAATGGC TTTTTTCTC   19

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 3974 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: both
       (D) TOPOLOGY: both

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
GGTACCTAAG TGAGTAGGGC GTCCGATCGA CGGACGCCTT TTTTTTGAAT TCGTAATCAT        60

GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG       120

CCGGAAGCAT AAAGTGTAAA GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG       180

CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA       240

TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG CTCTTCCGCT TCCTCGCTCA       300

CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG       360

TAATACGGTT ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC       420

AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC       480

CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC       540

TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC       600

TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA       660

GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC       720

ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA       780

ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG       840

CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA       900

GAAGAACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG       960

GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC      1020

AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT      1080

CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCGTCGA      1140

CAATTCGCGC GCGAAGGCGA AGCGGCATGC ATTTACGTTG ACACCATCGA ATGGTGCAAA      1200

ACCTTTCGCG GTATGGCATG ATAGCGCCCG GAAGAGAGTC AATTCAGGGT GGTGAATGTG      1260

AAACCAGTAA CGTTATACGA TGTCGCAGAG TATGCCGGTG TCTCTTATCA GACCGTTTCC      1320

CGCGTGGTGA ACCAGGCCAG CCACGTTTCT GCGAAAACGC GGGAAAAAGT GGAAGCGGCG      1380

ATGGCGGAGC TGAATTACAT TCCCAACCGC GTGGCACAAC AACTGGCGGG CAAACAGTCG      1440

TTGCTGATTG GCGTTGCCAC CTCCAGTCTG GCCCTGCACG CGCCGTCGCA AATTGTCGCG      1500

GCGATTAAAT CTCGCGCCGA TCAACTGGGT GCCAGCGTGG TGGTGTCGAT GGTAGAACGA      1560

AGCGGCGTCG AAGCCTGTAA AGCGGCGGTG CACAATCTTC TCGCGCAACG CGTCAGTGGG      1620

CTGATCATTA ACTATCCGCT GGATGACCAG GATGCCATTG CTGTGGAAGC TGCCTGCACT      1680

AATGTTCCGG CGTTATTTCT TGATGTCTCT GACCAGACAC CCATCAACAG TATTATTTTC      1740

TCCCATGAAG ACGGTACGCG ACTGGGCGTG GAGCATCTGG TCGCATTGGG TCACCAGCAA      1800

ATCGCGCTGT TAGCGGGCCC ATTAAGTTCT GTCTCGGCGC GTCTGCGTCT GGCTGGCTGG      1860
```

```
CATAAATATC TCACTCGCAA TCAAATTCAG CCGATAGCGG AACGGGAAGG CGACTGGAGT      1920

GCCATGTCCG GTTTTCAACA AACCATGCAA ATGCTGAATG AGGGCATCGT TCCCACTGCG      1980

ATGCTGGTTG CCAACGATCA GATGGCGCTG GGCGCAATGC GCGCCATTAC CGAGTCCGGG      2040

CTGCGCGTTG GTGCGGATAT CTCGGTAGTG GGATACGACG ATACCGAAGA CAGCTCATGT      2100

TATATCCCGC CGTTAACCAC CATCAAACAG GATTTTCGCC TGCTGGGGCA AACCAGCGTG      2160

GACCGCTTGC TGCAACTCTC TCAGGGCCAG GCGGTGAAGG CAATCAGCT GTTGCCCGTC       2220

TCACTGGTGA AAAGAAAAAC CACCCTGGCG CCCAATACGC AAACCGCCTC TCCCCGCGCG      2280

TTGGCCGATT CATTAATGCA GCTGGCACGA CAGGTTTCCC GACTGGAAAG CGGGCAGTGA      2340

GCGCAACGCA ATTAATGTAA GTTAGCGCGA ATTGTCGACC AAAGCGGCCA TCGTGCCTCC      2400

CCACTCCTGC AGTTCGGGGG CATGGATGCG CGGATAGCCG CTGCTGGTTT CCTGGATGCC      2460

GACGGATTTG CACTGCCGGT AGAACTCCGC GAGGTCGTCC AGCCTCAGGC AGCAGCTGAA      2520

CCAACTCGCG AGGGGATCGA GCCCGGGGTG GGCGAAGAAC TCCAGCATGA GATCCCCGCG      2580

CTGGAGGATC ATCCAGCCGG CGTCCCGGAA AACGATTCCG AAGCCCAACC TTTCATAGAA      2640

GGCGGCGGTG GAATCGAAAT CTCGTGATGG CAGGTTGGGC GTCGCTTGGT CGGTCATTTC      2700

GAACCCCAGA GTCCCGCTCA GAAGAACTCG TCAAGAAGGC GATAGAAGGC GATGCGCTGC      2760

GAATCGGGAG CGGCGATACC GTAAAGCACG AGGAAGCGGT CAGCCCATTC GCCGCCAAGC      2820

TCTTCAGCAA TATCACGGGT AGCCAACGCT ATGTCCTGAT AGCGGTCCGC CACACCCAGC      2880

CGGCCACAGT CGATGAATCC AGAAAAGCGG CCATTTTCCA CCATGATATT CGGCAAGCAG      2940

GCATCGCCAT GGGTCACGAC GAGATCCTCG CCGTCGGGCA TGCGCGCCTT GAGCCTGGCG      3000

AACAGTTCGG CTGGCGCGAG CCCCTGATGC TCTTCGTCCA GATCATCCTG ATCGACAAGA      3060

CCGGCTTCCA TCCGAGTACG TGCTCGCTCG ATGCGATGTT TCGCTTGGTG GTCGAATGGG      3120

CAGGTAGCCG GATCAAGCGT ATGCAGCCGC CGCATTGCAT CAGCCATGAT GGATACTTTC      3180

TCGGCAGGAG CAAGGTGAGA TGACAGGAGA TCCTGCCCCG GCACTTCGCC CAATAGCAGC      3240

CAGTCCCTTC CCGCTTCAGT GACAACGTCG AGCACAGCTG CGCAAGGAAC GCCCGTCGTG      3300

GCCAGCCACG ATAGCCGCGC TGCCTCGTCC TGCAGTTCAT TCAGGGCACC GGACAGGTCG      3360

GTCTTGACAA AAAGAACCGG GCGCCCCTGC GCTGACAGCC GGAACACGGC GGCATCAGAG      3420

CAGCCGATTG TCTGTTGTGC CCAGTCATAG CCGAATAGCC TCTCCACCCA AGCGGCCGGA      3480

GAACCTGCGT GCAATCCATC TTGTTCAATC ATGCGAAACG ATCCTCATCC TGTCTCTTGA      3540

TCAGATCTTG ATCCCCTGCG CCATCAGATC CTTGGCGGCA AGAAAGCCAT CCAGTTTACT      3600

TTGCAGGGCT TCCCAACCTT ACCAGAGGGC GCCCCAGCTG GCAATTCCGG TTCGCTTGCT      3660

GTCCATAAAA CCGCCCAGTC TAGCTATCGC CATGTAAGCC CACTGCAAGC TACCTGCTTT      3720
```

CTCTTTGCGC TTGCGTTTTC CCTTGTCCAG ATAGCCCAGT AGCTGACATT CATCCGGGGT 3780

CAGCACCGTT TCTGCGGACT GGCTTTCTAC GTGTTCCGCT TCCTTTAGCA GCCCTTGCGC 3840

CCTGAGTGCT TGCGGCAGCG TGAAGCTTAA AAAACTGCAA AAAATAGTTT GACTTGTGAG 3900

CGGATAACAA TTAAGATGTA CCCAATTGTG AGCGGATAAC AATTTCACAC ATTAAAGAGG 3960

AGAAATTACA TATG 3974

(2) INFORMATION FOR SEQ ID NO:26:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 112 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: both

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AAGCTTAAAA AACTGCAAAA AATAGTTTGA CTTGTGAGCG GATAACAATT AAGATGTACC 60

CAATTGTGAG CGGATAACAA TTTCACACAT TAAAGAGGAG AAATTACATA TG 112

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 415 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCTCGAGCTT TCCACGCAAA TCTCCAATAG CGTTGGTGTG CACACTCTTG CTCCCTGACC 60

TTGGCCCCTT GTGCTTTTCA TAGAGACACC CTGGGGAAGG CAGTTTTTGA TCATAATGGT 120

TGTCACCTGA GCTGGCATTA CCAAAAAGCT GGGCTCCACT GCCCACGCAG TACAGGTATG 180

GGTAGTAAGG ATAGCTGTTA TACACGGAGA CCAGGGAGGC CCGGGAGAAG GCCTCTCTTT 240

CAGGGCCGGC AAAGAGGAGT GCTTCTCCAA GTCTCCCAGC TCCGAGGAGA GCTGCTTTCG 300

CTTAGTCTTA TAGCGTCTGT TCTGGAACCA TATCTTCACT TGGGTCTCCG TGAGCTTGAG 360

GTTCTTGGCC AGGTGGGCCG TTCAGGGCGA TCAGGTACTT CTGATGGCTG AACTT 415

(2) INFORMATION FOR SEQ ID NO:28:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 507 base pairs

(B) TYPE: nucleic acid
          (C) STRANDEDNESS: both
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTTTTTTCCC GTCCGACCCT CCGGAGTTTA TTCACTTCCA GCGGTACTGG GCTGGACCGA   60

GCCTTGGTGG CATCCGATTC AGCTCCAGCA GCATCCCAGG TCCTATCCAG CATGGGCGGG   120

GCACCGCTGG TCAGAGCTCT GGCCTTCAGG AATCTTCGGA CTCGTCCCTT TCCTCGAAGT   180

CGGGCTCGGG CTCTGGGTCC GGTTCTGGCT CCAGTTCCGG CTCTGGTTCT GCTTCAGCCT   240

CGGGTTCCAG CTCTGGCTCT GGCTCCGGCT CCGGTTCAGC CTCTTTAGAG GCCTCAGCTT   300

CCAGCTCTGG CTCTGGCTCC TCCGGGGTCC CGGATTGCAC CGCCTCGGAG CGCTCAGGCC   360

CTCCGGCTAC GTCAGGGTCT GAAGCTCTGG GCAACCCGGG GCATGTGGTC ACCGAGCCGG   420

GTGCCTCCGG GTCCAGTGGC CTGGGCATGG AGGGTCATAG CTGCGGTCGC GGTAGCCCGG   480

ACCCACGTAC TGCGAGTCGA AAGCGGG   507

(2) INFORMATION FOR SEQ ID NO:29:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 533 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: both
          (D) TOPOLOGY: both

     (ii) MOLECULE TYPE: cDNA

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

NGCACAGCAG NNAGGCACTT GGAGGTCTTA TCTGTTNGAC TTGNTNAACA CTTCAGGCGC   60

CCTTCCAAGG CTTCCCCAAA CCCCTAAGCA GCCGCANAAG CGCTCCCGAG CTGCCTTCTC   120

CCACACTCAG GTGATCGAGT TGGAGAGGAA GTTCAGCCAT CAGAAGTACC TGTCGGCCCC   180

TGGAACGGGC CCACCTGGCC AAGAACCTCA AGCTCACGGA GACCCAAGTG AAGATATGGT   240

TCCAGAACAG ACGCTATAAG ACTAAGCGAA ACAGCTCTCC TCGGANCTGG GGAGACTTGG   300

AGAAGCACTC CTCTTTGNCC GGCCCTGAAN GAGGAGGCCN TCNCCCGGGC CNNCCTGGTC   360

NCCGTGTAAT AACAGCGANC CTTNANTACC CATACNTGTA ANGCNTGGGG CANNGGGAAG   420

CCCAGTTTNT TGGTAAANGG CCAGCTCCAG GTGAACAACC ATTAAGGATC AAAANGGCCT   480

TNCCCCAGGG TGTTNCATGG AAAAGCACAA GGGGCAAGGT CAGGAGCAAA AGN   533

(2) INFORMATION FOR SEQ ID NO:30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 318 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: both

   (ii) MOLECULE TYPE: cDNA

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
GGGAGCTGGA GCCAGAGCCA GAGCCAGATC TCAGGAAACT GGGGTCATTG CATAGAGGCT        60

GCCAGACAGT CTGCAGAGCT CAGCGGCCTG GGTTCAAACC TTCTCGCACA CTGCCACTGT       120

CGGTTACTTT GGCTTTCTAG AGCCAGATTC CTTGGCCATG AAATGGGTAC TGCTTACTTC       180

CCAGGTTATT TTGAGAATGA AGTGAGATGA AGTCAACAGT AGATGTATCT GTCCGTTGTC       240

CCTGCCCTGC TGTGGGGATG ACAGAGTGAT TTTGGACAAG ACCCAAGGNC TCGCTGGGCA       300

TCACTGGTCT TTCTTCAG                                                     318
```